

# Efficient Simulations of Individual Based Models for Adaptive Dynamics and the Canonical Equation

Vaibhav Madhok<sup>1,\*</sup>

<sup>1</sup>*Department of Zoology, University of British Columbia,  
6270 University Boulevard, Vancouver B.C. Canada, V6T 1Z4*

We propose a faster algorithm for individual based simulations for adaptive dynamics based on a simple modification to the standard Gillespie Algorithm for simulating stochastic birth death processes. We provide an analytical explanation that shows that simulations based on the modified algorithm, in the deterministic limit, lead to the same equations of adaptive dynamics as well as same conditions for evolutionary branching as those obtained from the standard Gillespie algorithm. Based on this algorithm, we provide an intuitive and simple interpretation of the canonical equation of adaptive dynamics. With the help of examples we compare the performance of this algorithm to the standard Gillespie algorithm and demonstrate its efficiency. We also study an example using this algorithm to study evolutionary dynamics in a multi-dimensional phenotypic space and study the question of predictability of evolution.

## I. INTRODUCTION

Evolution is essentially a birth death process involving mutations and selection. According to the traditional view, evolution optimizes scalar phenotypes like beak size, fecundity, body size etc. The dictum, “Survival of the fittest”, leads to the fittest type “winning”. Evolution is thus viewed as a dynamical system converging to an equilibrium or a steady state. Such a view fails to take into account interactions among the individuals comprising a population or interactions among various individuals present in a geographical location. The birth and death rates of the individual depend on its phenotypic composition as well as its interaction with the “environment”. The “environment” consists of both the “abiotic” environment like temperature, climate etc. and the “biotic” environment which can be, for example, the types of all other individuals with which it interacts, for example, the competition it faces with regards to resources such as food and shelter. Another example of the “biotic” environment is the presence of “prey” or “predator” type of individuals in the population [1]. Studying evolution involves modeling this birth death process in order to address questions about origin of diversity, speciation and predictability of evolution. In order to have such a view, we need to take into account certain ecological mechanisms like the interaction among organisms present in the same geographical location - a phenomenon known as frequency dependence [1–4]. Adaptive dynamics is a framework that incorporates frequency dependent selection originating from ecological interactions and competition among species [1]. Under this framework, evolution is regarded as a continuous trajectory in the phenotypic space that unravels as rare mutants with a higher “fitness” invade a resident population. The evolutionary dynamics in the above framework is described by the “canonical equation” [5], which, for one

dimension, is given by

$$\frac{dx}{dt} = \gamma \frac{\partial f(\mathbf{x}, \mathbf{y})}{\partial y} \Big|_{\mathbf{y}=\mathbf{x}}. \quad (1)$$

Here  $\gamma$  is a constant of proportionality that scales the rate of evolution. The function,  $f(\mathbf{x}, \mathbf{y})$ , is the measure of fitness of a rare mutant,  $y$ , in a monomorphic population with a resident trait  $x$  [5]. The derivative of the invasion fitness function,  $f(\mathbf{x}, \mathbf{y})$ , with respect to the mutant phenotype coordinate,  $y$ , gives the direction of motion of the evolutionary trajectory. Mutations cause the occurrence of rare types which can possibly invade the resident depending on their relative growth rate with respect to the resident which is given by the selection gradient (RHS of Eq 1).

The derivation of the canonical equation takes into account the stochastic nature of the birth-death process and is based on the ecological processes influencing the birth and death at the level of an individual [5–7]. Such a formulation assumes that mutations are small, they occur very rarely and the resident population is taken to be of infinite size [7]. The canonical equation (Eq. 1) describes the dynamics of a monomorphic population in the trait space driven by mutations, selection and invasion.

In order to address questions like the origin of diversity, to study the stochastic effects in evolution and to study the issue of predictability of evolution we need to construct individual based models. Individual based stochastic models based on the deterministic adaptive dynamics are a natural generalization of the latter. In these models, we represent the population as a “cloud” of points in the phenotypic space. Each individual, represented by a point in the cloud, has a given birth and death rate depending on the location of the point in the phenotypic space. Individuals with a higher birth rate have a greater probability to reproduce and the individuals with a higher death rate have a greater probability to die.

Individual-based realizations of the model are typically based on the Gillespie algorithm [8]. Often, simulations of these models take a long time for one to see interesting effects like diversification in the phenotype space. Moreover, once diversification occurs, especially in high dimensional phenotypic spaces, it usually leads to an increase in population of

---

\*Electronic address: vmadhok@gmail.com

individuals and therefore making simulations even slower. In this work, we accomplish three things. First, we propose a simple modification of the standard Gillespie algorithm and make the individual based simulations more efficient. With the help of numerical simulations, we demonstrate how this leads to faster simulations. Secondly, using this algorithm and under certain assumptions, we give a very intuitive and simple interpretation of the canonical equation of adaptive dynamics. Thirdly, we discuss the issue of the predictability of evolution by considering individuals based models in high dimensional phenotypic spaces and demonstrate the “evolutionary butterfly effect”.

## II. INDIVIDUAL BASED MODELS TO MACROSCOPIC DETERMINISTIC ADAPTIVE DYNAMICS

We begin with the widely studied logistic model [9],

$$\frac{\partial N(\mathbf{x}, t)}{\partial t} = rN(\mathbf{x}, t) \left( 1 - \frac{\int \alpha(\mathbf{x}, \mathbf{y}) N(\mathbf{y}, t) d\mathbf{y}}{K(\mathbf{x})} \right). \quad (2)$$

We will call this the “canonical” model, as later on we will also be interested in a general family of such models.

Here  $N(\mathbf{x}, t)$  is the population density of the phenotype  $\mathbf{x}$  at time  $t$  and  $K(\mathbf{x})$  is the carrying capacity of the population consisting entirely of  $\mathbf{x}$  individuals. The competition between individuals of phenotypes  $\mathbf{x}$  and  $\mathbf{y}$  is given by the competition kernel  $\alpha(\mathbf{x}, \mathbf{y})$  and an individual of phenotype  $\mathbf{x}$  faces competition with an effective density  $\int \alpha(\mathbf{x}, \mathbf{y}) N(\mathbf{y}, t) d\mathbf{y}$ . We set the intrinsic growth parameter  $r$  and  $\alpha(\mathbf{x}, \mathbf{x})$  to be equal to 1. We consider  $\mathbf{x}$  and  $\mathbf{y}$  to be vectors of dimension  $d$  describing a multidimensional phenotypic space.

The canonical equations of adaptive dynamics can be derived from the stochastic individual based model in the limit of rare mutations, small mutational effects and infinite population sizes. Under these assumptions, Dieckmann and Law showed that adaptive dynamics is the first order approximation of the mean path averaged over infinitely many realizations of the stochastic simulations obtained from the individual based model. In order to derive the adaptive dynamics of trait  $\mathbf{x}$ , we assume a monomorphic (Dirac-delta distribution) resident population in trait  $\mathbf{x}$  with a globally stable equilibrium density given by  $K(\mathbf{x})$  independent of the dimension  $d$  of  $\mathbf{x}$ . A key ingredient to obtain the ODEs describing the adaptive dynamics is the *invasion fitness* function whose gradient determines the direction of selection forces and hence of the deterministic evolutionary trajectory. The *invasion fitness* function is the difference between the per capita birth and per capita death rates of a rare mutant in a monomorphic population. The *invasion fitness*, for the logistic map, of a rare mutant  $\mathbf{y}$  is its per capital rate of growth in a resident population with phenotype  $\mathbf{x}$  and is given by

$$f(\mathbf{x}, \mathbf{y}) = 1 - \frac{\alpha(\mathbf{x}, \mathbf{y}) K(\mathbf{x})}{K(\mathbf{y})}. \quad (3)$$

In the above example for the logistic equation, the selection due to frequency dependence influences the death rates while

the birth rates are neutral. In general, birth rates might also depend on interactions between individuals and the fitness function will have a more general form involving competition kernels for both birth and death.

From the *invasion fitness*, we can derive the selection gradient to be

$$s_i(\mathbf{x}) = \frac{\partial f(\mathbf{x}, \mathbf{y})}{\partial y_i} \Big|_{\mathbf{y}=\mathbf{x}} = - \frac{\partial \alpha(\mathbf{x}, \mathbf{y})}{\partial y_i} \Big|_{\mathbf{y}=\mathbf{x}} + \frac{\partial K(\mathbf{x})}{\partial x_i} \frac{1}{K(\mathbf{x})}. \quad (4)$$

Finally, the canonical equation of adaptive dynamics in terms of the selection gradient is given by

$$\frac{d\mathbf{x}}{dt} = M(\mathbf{x}) \cdot s(\mathbf{x}). \quad (5)$$

Equation (5) is a system of  $d$  dimensional ODEs that describe a trajectory that may converge to a fixed point, a limit cycle, a quasi periodic orbit or exhibit chaos depending on the nature of the selection gradient  $s(x)$  and the initial conditions. Here  $M(x)$  describes the mutational process and for simplicity we assume it to be a  $d \times d$  identity matrix.

Individual based models describe the interactions between discrete individuals that possess multi-dimensional adaptive traits. Such a microscopic description is more general and in a sense more fundamental as compared to its deterministic macroscopic approximation like the adaptive dynamics which can be derived from it under suitable assumptions. In order to simulate individual based models, we adapt the Gillespie algorithm [8] for simulating chemical reactions to population biological stochastic processes. Individuals are treated as particles, and birth and death events as chemical reactions. In the deterministic logistic model, the birth rate of phenotype  $i$ ,  $\rho_i$ , with trait value  $\mathbf{x}_i$  is 1. When a birth event happens, a new individual is added to the population with a phenotype that is offset from the parent by a small mutation, which happens with probability  $\mu$ , chosen from a uniform distribution with a small amplitude. Throughout this paper we choose  $\mu = 1.0$ . Its death rate,  $\delta_i$  is given by  $N_{eff}(i)/K(\mathbf{x}_i)$ , where  $N_{eff}$  is the effective density experienced by phenotype  $\mathbf{x}$ . For the individual based model, if there are  $N$  individuals,  $\mathbf{x}_1, \dots, \mathbf{x}_N$ , at time  $t$  then the effective density experienced by individual  $i$  is

$$N_{eff}(i) = \sum_{j \neq i} \alpha(\mathbf{x}_j, \mathbf{x}_i), \quad (6)$$

and, therefore, the death rate is given by

$$N_{eff}(i) = \sum_{j \neq i} N_{eff}/K(\mathbf{x}_i). \quad (7)$$

The individual based simulations are then implemented stochastically in the familiar way by performing one birth or death event [8] that constitutes one computational step that advances the system from time  $t$  to  $t + \Delta t$ . In the above example for the logistic equation, the selection due to frequency dependence influences the death rates while the birth rates are neutral. In general, both the birth as well as the death rates can be affected by frequency dependent interactions among individuals described by competition terms like  $\alpha_{birth}(\mathbf{x}_j, \mathbf{x}_i)$  and  $\alpha_{death}(\mathbf{x}_j, \mathbf{x}_i)$ .

### III. A MODIFIED ALGORITHM FOR INDIVIDUAL BASED SIMULATIONS

As mentioned above, individual-based realizations of the model are based on the Gillespie algorithm [8] as described above and consists of assigning each individual  $i$  a constant reproduction rate  $\rho_i = 1$  and a death rate  $\delta_i = \sum_{j \neq i} \alpha(\mathbf{x}_i, \mathbf{x}_j)/K(\mathbf{x}_i)$ , as defined by logistic ecological dynamics. The total rate is given by the sum of all individual rates  $U = \sum_i (\rho_i + \delta_i)$ . A particular “event”, birth or death, is chosen at random with probability equal to the rate of this event divided by the total probability rate  $U$ . If a birth event is chosen, a new individual is added to the population with a phenotype that is offset from the parent by a small mutation chosen from a uniform distribution with a small amplitude. Modified Gillespie algorithm consists of simply choosing the individual with the maximum death rate and eliminating it when a death event is chosen. That is, we select  $\max(\delta_1, \delta_2, \dots, \delta_n)$  and eliminate that individual. If both the birth as well as the death rates are affected by frequency dependent interactions among individuals described by competition terms like  $\alpha_{birth}(\mathbf{x}_j, \mathbf{x}_i)$  and  $\alpha_{death}(\mathbf{x}_j, \mathbf{x}_i)$ , when the birth event is chosen, we select the individual with the maximum birth rate,  $\rho_i$ , and a new individual is added to the population with a phenotype that is offset from the parent by a small mutation chosen from a uniform distribution with a small amplitude. We treat the death events analogously.

It turns out that this simple modification has remarkable consequences as far as simulations are concerned. We now discuss the consequences of this modification on the efficiency of the algorithm and also on the adaptive dynamics it gener-

ates.

#### A. Analysis of the Algorithm

In order to understand the above algorithm and the reason behind its speed, we first consider a logistic model of the form

$$\frac{\partial N(\mathbf{x}, t)}{\partial t} = N(\mathbf{x}, t) \left( 1 - \left( \frac{\int \alpha(\mathbf{x}, \mathbf{y}) N(\mathbf{y}, t) d\mathbf{y}}{K(\mathbf{x})} \right)^m \right). \quad (8)$$

The above model reduces to Eq. 2 for  $m = 1$ . Assuming the resident to be at the equilibrium population density, given by the Dirac delta function,  $N(x') = K(x)\delta(x - x')$ . Then the per capita growth of a mutant  $y$  is given by

$$\frac{\partial N(\mathbf{y}, t)}{\partial t N(\mathbf{y}, t)} = \left( 1 - \left( \frac{\int \alpha(\mathbf{x}', \mathbf{y}) K(x)\delta(x - x') dx'}{K(\mathbf{y})} \right)^m \right). \quad (9)$$

Therefore, the invasion fitness, as defined above is given by

$$f(\mathbf{x}, \mathbf{y}) = \left( 1 - \left( \frac{\alpha(\mathbf{x}, \mathbf{y}) K(x)}{K(\mathbf{y})} \right)^m \right). \quad (10)$$

From this, we can derive the selection gradient to be

$$s_i(\mathbf{x}) = \frac{\partial f(\mathbf{x}, \mathbf{y})}{\partial y_i} \Big|_{\mathbf{y}=\mathbf{x}} = m \left( \frac{\alpha(\mathbf{x}, \mathbf{y}) K(x)}{K(\mathbf{y})} \right) \Big|_{\mathbf{y}=\mathbf{x}} \left( - \frac{\partial \alpha(\mathbf{x}, \mathbf{y})}{\partial y_i} \Big|_{\mathbf{y}=\mathbf{x}} + \frac{\partial K(\mathbf{x})}{\partial x_i} \frac{1}{K(\mathbf{x})} \right), \quad (11)$$

which simplifies to

$$s_i(\mathbf{x}) = \frac{\partial f(\mathbf{x}, \mathbf{y})}{\partial y_i} \Big|_{\mathbf{y}=\mathbf{x}} = -m \left( \frac{\partial \alpha(\mathbf{x}, \mathbf{y})}{\partial y_i} \Big|_{\mathbf{y}=\mathbf{x}} + \frac{\partial K(\mathbf{x})}{\partial x_i} \frac{1}{K(\mathbf{x})} \right). \quad (12)$$

Therefore, the above selection gradient is  $m$  times the selection gradient given in Eq. 4.

Finally, the canonical equation of adaptive dynamics in terms of the selection gradient is given by

$$\frac{d\mathbf{x}}{dt} = m M(\mathbf{x}) \cdot s(\mathbf{x}), \quad (13)$$

which gives the same set of ODEs up to the scaling constant  $m$  as obtained previously with the canonical logistic model. We also note that the Eq.13 evolves  $m$  times faster than Eq. 5.

In constructing the individual based models for the logistic model in Eq. 8, we note that the birth rates remain constant at  $\rho_i = 1$ , while the death rate for the individual  $i$  is given by  $\delta_i = \left( \sum_{j \neq i} \alpha(\mathbf{x}_i, \mathbf{x}_j) / K(\mathbf{x}_i) \right)^m$ . Therefore, the vector of these death rates can be expressed as  $(\delta_1^m, \delta_2^m, \dots, \delta_n^m)$  where

$(\delta_1, \delta_2, \dots, \delta_n)$  are the death rates of individuals corresponding to the “canonical” map. Let  $k$  be the individual with the highest death rate. This means  $\max(\delta_1^m, \delta_2^m, \dots, \delta_n^m) = \delta_k^m$ . When  $m$  is sufficiently large, we have  $\delta_k^m \gg \delta_i^m$  for all  $i \neq k$ . Therefore the probability of choosing  $k$  to die is close to 1. But this is exactly what we do in our modified algorithm! We deterministically pick the individual with the highest death rate and eliminate it. Therefore, the modified algorithm can be considered to be an individual based simulation of a birth death process governed by the logistic model given by Eq. 8. The value of  $m$  in Eq. 8 should be sufficiently large such that, we satisfy,  $\delta_k^m \gg \delta_i^m$  for all  $i \neq k$ . We have already shown that the individual based simulations based on this model, in the deterministic limit, give rise to exactly the same equations of adaptive dynamics as the “canonical” model up to the scaling constant  $m$ .

### B. Conditions for Diversification

We have shown that the individual based simulations based on the modified algorithm gives us the same adaptive dynamics equations as the standard algorithm. ODEs describing the adaptive dynamics are first order differential equations describing the evolution of a monomorphic population in the phenotypic space. In order to study diversification, we need

to explore higher orders of the fitness function.

Adaptive diversification has been extensively studied at singular points of the dynamics that also happen to be evolutionarily unstable [1, 4, 10]. The canonical equation of adaptive dynamics has the form

$$\frac{dx_i}{dt} = g_i(x), \quad (14)$$

where the functions  $g_i$  are given as

$$\begin{pmatrix} g_1(x) \\ \vdots \\ g_d(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial f(x,y)}{\partial y_1} \big|_{y=x} \\ \vdots \\ \frac{\partial f(x,y)}{\partial y_d} \big|_{y=x} \end{pmatrix}. \quad (15)$$

Singular points of the dynamics are the points,  $\mathbf{x}^*$  at which the R.H.S. of Eq. (14) is zero. Singular point,  $\mathbf{x}^*$  is called *convergence stable* if the dynamics starting at all points sufficiently close to it converge to  $\mathbf{x}^*$  eventually. This occurs when the Jacobian at the singular point,  $J(\mathbf{x}^*) < 0$ . These points are of great importance in adaptive dynamics as they can be potentially the points at which diversification occurs.

For scalar traits, this can be seen by Taylor expanding the invasion fitness function with respect to the mutant  $y$  to the second order, we get

$$f(x, y) = f(x, x) + \frac{\partial f(x, y)}{\partial y} \big|_{y=x} \cdot (y - x) + \frac{\partial^2 f(x, y)}{\partial y^2} \big|_{y=x} \cdot \frac{(y - x)^2}{2}. \quad (16)$$

The first term on the RHS,  $f(x, x)$  is zero for all  $x$ . Usually, the evolutionary dynamics can be accurately described by the first order term in (16) or, in other words, by the canonical equation of adaptive dynamics and we do not need to consider higher orders terms. It is only in the neighborhood of singular points, we have  $\frac{\partial f(x, y)}{\partial y} \big|_{y=x} \rightarrow 0$ , the second order term in Eq. (16) becomes significant. In particular, at the singular point we have  $\frac{\partial f(x, y)}{\partial y} \big|_{y=x} = 0$  and if  $\frac{\partial^2 f(x, y)}{\partial y^2} \big|_{y=x} < 0$ , no nearby mutants can invade the resident population that is monomorphic and we have the conditions for evolutionary stability. In contrast,  $\frac{\partial^2 f(x, y)}{\partial y^2} \big|_{y=x} > 0$  is the condition for

evolutionary instability, or potential evolutionary branching points, as the mutant now can potentially invade the resident.

It is easy to see that the fitness functions derived from Eq. 2 and Eq. 8 yield the same conditions for diversification. For the Eq. 2 we have

$$\frac{\partial^2 f(x, y)}{\partial y^2} \big|_{y=x} = -\frac{\partial}{\partial y} \left( \frac{\partial}{\partial y} \left( \frac{\alpha(x, y)k(x)}{k(y)} \right) \right) \bigg|_{y=x}, \quad (17)$$

and for the invasion fitness function in Eq. 10, the second derivative is

$$\frac{\partial^2 f(x, y)}{\partial y^2} \big|_{y=x} = -m \left( \frac{\partial}{\partial y} \left( \frac{\alpha(x, y)k(x)}{k(y)} \right) \right) \bigg|_{y=x}^2 - m \frac{\partial}{\partial y} \left( \frac{\partial}{\partial y} \left( \frac{\alpha(x, y)k(x)}{k(y)} \right) \right) \bigg|_{y=x}. \quad (18)$$

At equilibrium, when we have  $\frac{\partial f(x, y)}{\partial y} \big|_{y=x} = 0$ , the first term

on the RHS is zero. And the second term coincides with Eq.

17. Therefore, the conditions for potential diversification are identical for both maps. Hence we conclude that to study diversification through the modified algorithm, we can expect to get similar results as we would using the standard Gillespie algorithm.

#### IV. INTERPRETATION OF THE CANONICAL EQUATION

Canonical equation can be viewed as dynamics that unravel as individuals go through the birth-death process described above. We will give a geometric interpretation of the canonical equation through the birth-death process.

We first consider the time reversal of the dynamics in the logistic model. We first recall that a birth-death process in the logistic model has birth and death rates given by  $\rho_i = 1$  and  $\delta_i = \sum_{j \neq i} \alpha(\mathbf{x}_i, \mathbf{x}_j)/K(\mathbf{x}_i)$ . If we simply interchange the birth rates with the death rates, i.e.,  $\delta_i = 1$  and  $\rho_i = \sum_{j \neq i} \alpha(\mathbf{x}_i, \mathbf{x}_j)/K(\mathbf{x}_i)$ , this gives us the time reversal of the evolutionary dynamics given by the logistic model. This can also be seen from the expression for the invasion fitness, Eq. 3 where the first term on the RHS represents the per capita birth rate and the second term represents the per capita death rate [5]. Interchanging birth and death rates merely changes the sign of the fitness function and the selection gradient thereby reversing the evolutionary dynamics given in Eq. 5. Therefore, a birth event in the forward time is equivalent to a death event in the time reversed picture and, likewise, a death event in the forward time is equivalent to a birth event in the time reversed picture.

When a birth event is chosen, a new individual is added to the population with a phenotype that is offset from the parent by a small mutation chosen from a uniform distribution with a small amplitude. Therefore, one can view the birth event as the displacement of the population cloud by a very small amount. In the time reversed picture, this birth event becomes a death event and will lead to the displacement of the population cloud in the opposite direction.

Considering the logistic model, we consider a time period  $\Delta\tau$  which is very small compared to the total time in which the dynamics unravels but nonetheless large enough that we have a large number of birth and death events occurring. For infinite population sizes, we will have a large number of birth-death events happening even when  $\Delta\tau$  becomes very small. We also assume, the number of birth events  $N_b$  will be equal to the number of death events  $N_d$  in the time period  $\Delta\tau$ . This is because the population is assumed to be at equilibrium and, the probability of birth is equal to the probability of death event, i.e  $P(\text{Birth}) = P(\text{Death})$ . Therefore, invoking the law of larger numbers, the sequence of birth-death events will be typical sequences with  $N(\text{birth}) = N(\text{death})$ . Consider a monomorphic population cloud at a particular point  $\mathbf{x}$  in the phenotypic space. In the limit of very small mutations and a sufficiently small  $\Delta\tau$ , a single birth event will cause a displacement of the population cloud given by  $\Delta r_i$ . Since for the logistic model, the birth events are neutral, and therefore  $\sum_i \Delta r_i = 0$ . The displacement by a single death event is,  $\Delta q$ , and the direction of displacement is in the opposite direction

to the dying individual  $k$ , where  $\delta_k = \max(\delta_1, \delta_2, \dots, \delta_n)$ .  $\delta_k = \max(\sum_{j \neq i} \alpha(\mathbf{x}_i, \mathbf{x}_j)/K(\mathbf{x}_i))$ . Assuming all individuals except  $k$  to be a monomorphic population,  $\delta_k = \max(\alpha(\mathbf{x}, \mathbf{x} + \Delta q)/K(\mathbf{x} + \Delta q))$ . Thus, we seek to determine a vector displacement  $\Delta q$  that will maximize  $\delta_k$ . If  $\Delta q$  is in the direction of the gradient of the function  $\frac{\alpha(\mathbf{x}, \mathbf{y})K(\mathbf{x})}{K(\mathbf{y})}$  with respect to  $y$  at the value  $x$ . We assume the same direction for  $\Delta q$  for all the death events occurring in the time interval  $\Delta\tau$ . Therefore, the total net displacement is given by

$$\sum_i \Delta r_i + \sum_j \Delta q_j = N_{\text{death}} \nabla_{\mathbf{y}|\mathbf{y}=\mathbf{x}} f(\mathbf{x}, \mathbf{y}) \cdot |\Delta q|. \quad (19)$$

Taking the limits,  $N_{\text{death}} \rightarrow \infty$ ,  $|\Delta q| \rightarrow 0$  and  $\tau \rightarrow 0$  we get the canonical equation of adaptive dynamics,

$$\frac{d\mathbf{x}}{dt} = \gamma \nabla_{\mathbf{y}|\mathbf{y}=\mathbf{x}} f(\mathbf{x}, \mathbf{y}). \quad (20)$$

In general and as mentioned above, birth rates might also depend on interactions between individuals and the fitness function will have a more general form involving competition kernels for both birth and death. Let,  $\alpha_{\text{birth}}(\mathbf{x}_j, \mathbf{x}_i)$  and  $\alpha_{\text{death}}(\mathbf{x}_j, \mathbf{x}_i)$  be the competition kernels for birth and death respectively. Since we assume an infinite population limit at equilibrium we have  $N(\text{birth}) = N(\text{death})$ . In this case, we will need to add up the displacement by both birth as well as death events. For a single birth event, the displacement is,  $\Delta q_b$ , and the direction of displacement is in the direction of the newly born individual  $k$ , where  $\rho_k = \max(\rho_1, \rho_2, \dots, \rho_n)$ .  $\rho_k = \max(\sum_{j \neq i} \alpha_{\text{birth}}(\mathbf{x}_i, \mathbf{x}_j)/K(\mathbf{x}_i))$ . Assuming all individuals, except  $k$ , as a monomorphic population,  $\rho_k = \max(\alpha_{\text{birth}}(\mathbf{x}, \mathbf{x} + \Delta q_b)/K(\mathbf{x} + \Delta q_b))$ . Thus, we seek to determine a vector displacement  $\Delta q_b$  that will maximize  $\rho_k$ . If  $\Delta q_b$  is in the direction of the gradient of the function  $\frac{\alpha_{\text{birth}}(\mathbf{x}, \mathbf{y})K(\mathbf{x})}{K(\mathbf{y})}$  with respect to  $y$  at the value  $x$ . We assume the same direction for  $\Delta q_b$  for all the birth events occurring in the time interval  $\Delta\tau$ . The calculation for the net displacement,  $\Delta q_d$ , due to death events follows similarly. Therefore, the total net displacement, with  $N(\text{birth}) = N(\text{death})$ , is given by

$$\sum_i \Delta r_i + \sum_j \Delta q_j = N_{\text{death}} \nabla_{\mathbf{y}|\mathbf{y}=\mathbf{x}} f'(\mathbf{x}, \mathbf{y}) \cdot |\Delta q|, \quad (21)$$

where,  $f'(\mathbf{x}, \mathbf{y}) = \left( \frac{\alpha_{\text{birth}}(\mathbf{x}, \mathbf{y})K(\mathbf{x})}{K(\mathbf{y})} - \frac{\alpha_{\text{death}}(\mathbf{x}, \mathbf{y})K(\mathbf{x})}{K(\mathbf{y})} \right)$  is the per capita growth rate of a mutant  $y$  in a resident population. In other words,  $f'(\mathbf{x}, \mathbf{y})$  is the invasion fitness function and we recover canonical equation of adaptive dynamics.

We interpret the above equation as the translation of a population cloud along the direction of instantaneous gradient of the fitness function with respect to the mutant phenotype. It is remarkable that in this derivation, unlike the standard derivation [5] which essentially considers the evolutionary trajectory as a sequence of invasions by the mutant phenotypes, we do not need to explicitly invoke concepts like ‘‘invasion fitness’’ but recover it through the translation of the population cloud undergoing a birth-death process.

Equation 21 also suggests a faster way to implement birth death process when we have competition kernels for both birth and death events. At each iteration, we simply choose exactly one birth and one death event. For the birth event, the individual with the maximum birth rate is chosen and a new individual in its vicinity is produced. For the death event, as before, the individual with the maximum death rate is chosen to die. It is clear from the form of the invasion fitness function, Eq. 3, consisting of two separate terms for birth and death that such an algorithm will satisfy the canonical equation of adaptive dynamics. The advantage we get is faster simulations and also the populations in the stochastic simulations does not increase and remains manageable. This is in contrast to the standard algorithm, where the population typically increases after evolutionary branching and the simulations computationally intensive after some time.

## V. EXAMPLES

In this section, we illustrate the efficiency of the modified algorithm with the help of numerical examples. We will also address the issue of unpredictability in evolution.

### A. Evolution of a single trait

Our first example compares the performance of the standard and modified Gillespie algorithm for a single trait. For this, we use the “canonical” logistic model given in Eq. 2. We consider a Gaussian competition kernel given by

$$\alpha(x, y) = \exp \left[ -\frac{(x - y)^2}{2\sigma_\alpha^2} \right]. \quad (22)$$

The magnitude of  $\sigma_\alpha$  tells us about the strength of the competition.

For the carrying capacity, we assume a Gaussian form as well,

$$K(x) = K_0 \exp \left[ -\frac{(x - x_0)^2}{2\sigma_k^2} \right]. \quad (23)$$

$K_0$  refers to the maximum carrying capacity and  $x_0$  is the phenotype with the maximum carrying capacity. The condition for diversification [1],  $\frac{\partial^2 f(x, y)}{\partial y^2} \Big|_{y=x} > 0$ , becomes

$$\sigma_\alpha < \sigma_k. \quad (24)$$

This can be interpreted as a competition between the relative effect of the competition and carrying capacity at the singular point. There has been great progress theoretically as well as a large amount of empirical evidence for sympatric speciation [11–33]. Fig 1 gives an example of the comparison between the standard implementation of the stochastic birth-death process and our modified procedure. The details of parameter values for the individual based simulations is given in the caption. We see the modified algorithm is far more efficient than the

standard Gillespie algorithm. While the latter takes roughly  $8 \times 10^4$  iterations for the first evolutionary branching to occur, the former shows multiple evolutionary branching events over  $2 \times 10^4$  iterations.

### B. Evolutionary chaotic dynamics in a multi dimensional phenotypic space and the question of predictability of evolution

We assume that the complexity of the epistatic interactions between phenotypic components is determined by the competition kernel  $\alpha(x, y) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ , which is, in general, a complicated non-linear function. In particular, when we have the competition kernel of the “canonical” form [33],

$$\alpha(\mathbf{x}, \mathbf{y}) = \exp \left( \sum_{j=1}^d b_{ij} x_j (x_i - y_i) + \sum_{j,k=1}^d a_{ijk} x_j x_k (x_i - y_i) \right), \quad (25)$$

we get

$$\frac{\partial \alpha(\mathbf{x}, \mathbf{y})}{\partial y_i} \Big|_{y=x} = - \sum_{j=1}^d b_{ij} x_j - \sum_{j,k=1}^d a_{ijk} x_j x_k. \quad (26)$$

For the carrying capacity function, we assume it to be of the form  $K(x) = \exp(-\sum_i x_i^4/4)$ . Then (5) can be written as

$$\frac{dx_i}{dt} = \sum_{j=1}^d b_{ij} x_j + \sum_{j,k=1}^d a_{ijk} x_j x_k - x_i^3, \quad i = 1, \dots, d. \quad (27)$$

The above set of ODEs describe the adaptive dynamics of the monomorphic population in the phenotypic space. In [34, 35], it was shown that the dynamical system described by Eq. (27) exhibits chaos with almost unit probability for sufficiently high dimensional phenotypic space, i.e.  $d \geq 50$ .

In particular, we are interested when (5) results in chaotic behavior. The above model can also be transformed to give the adaptive dynamics that coincides with the known chaotic systems like the Lorenz model. We just need to fix the coefficients  $a_i$ ,  $b_j$  and  $c_k$  of the competition kernel in the above model. For example, specific choice of coefficients will lead to Lorenz model and other popular models in the literature to explore the stochastic effects inherent in the birth death simulations when the corresponding equations for the adaptive dynamics are chaotic. Assuming the adaptive dynamics to be governed by the equations of the Lorenz model, we then proceed to study and construct the corresponding individual based model and simulate it the modified version of the standard algorithm.

Starting two simulations from exactly the same initial conditions and same mutation rates we see that the evolutionary trajectories follow completely different paths after sometime 2. Therefore the intuition about the evolutionary trajectories occupying two completely different regions of phenotypic space is correct. And this occurs purely due to a combination of stochastic effects due to mutations and the underlying

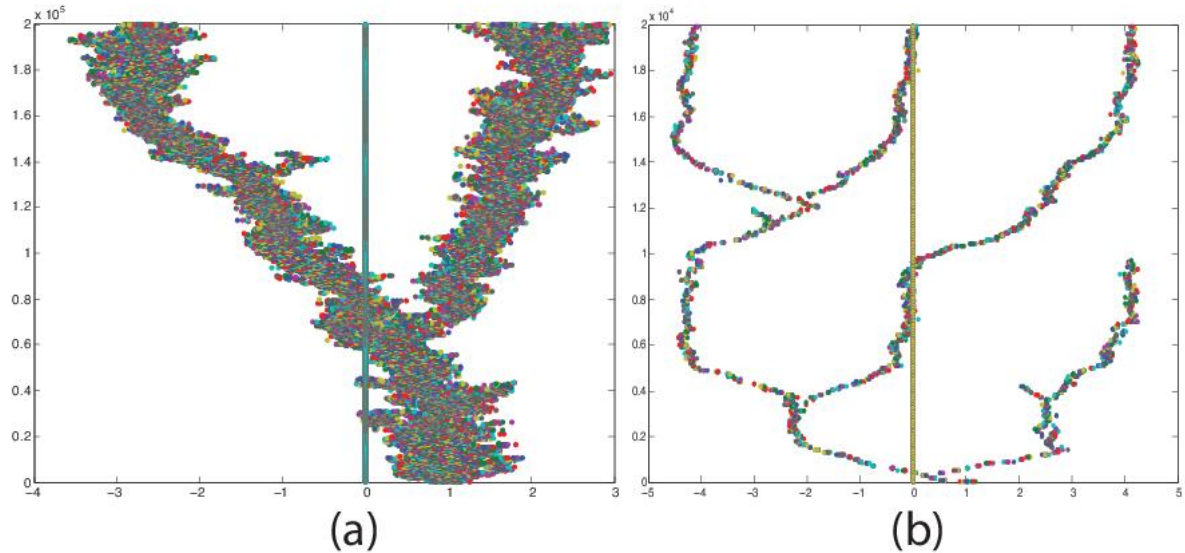


FIG. 1: Side by side comparison of the performance of simulation of single trait adaptive dynamics over the number of iterations on the computer (vertical axis) using the standard Gillespie algorithm on the left (Fig 1(a)) and the modified algorithm on the right (Fig 1(b)). Parameter values for both (a) and (b):  $\sigma_k = 2.0$ , a Gaussian competition kernel with  $\sigma_\alpha = 0.75$ .  $K_0 = 200$ . Initial population of 200 individuals was initialized about  $x = 1.0$  with a small variance. The mutation size after the birth event has a magnitude  $\sigma_\mu = 0.1$  and occurs with a probability  $\mu = 1.0$ . We see the modified algorithm is far more efficient than the standard Gillespie algorithm. While the latter takes roughly  $8 \times 10^4$  iterations for the first evolutionary branching to occur, the former shows multiple evolutionary branching events over  $2 \times 10^4$  iterations.

chaotic adaptive dynamics. Firstly, this tells us, contrary to the traditional view, that evolution is not a directional optimization process that goes to equilibration [34]. It might be a chaotic trajectory that is ergodic and sensitive to the initial conditions as well as the to the nature of very small mutational changes for exactly the same initial conditions. Even more fascinating is the sensitive dependence of a chaotic evolutionary trajectory on very small mutations, or the “butterfly effect” which casts a serious doubt over our ability to predict evolution over long time scales.

Fig 2 demonstrates the “evolutionary butterfly effect” for two instances of individual based simulations starting from exactly the same initial conditions when the underlying adaptive dynamics is chaotic. This is contrasted with individual based simulations of a cyclic adaptive dynamics which remains predictable throughout its course.

## VI. CONCLUSION

Individual based models are essential in order to study various aspects of evolution like diversification, predictability, adaptive radiation and frequency dependence. It has been observed that these models follow the adaptive dynamics trajectories remarkably well, even when the underlying adaptive dynamics is chaotic [33]. In this work, we derive the canonical equation of adaptive dynamics in a single step from a modified algorithm for simulating individual based models. We therefore provide a very simple and intuitive way of under-

standing the correspondence between adaptive dynamics and individual based models. Using a simple modification to the Gillespie algorithm, we were able to simulate individual based birth-death models with significant efficiency. With the help of examples, we demonstrate the efficiency of the modified algorithm. We also demonstrate how the modified algorithm can be interpreted as simulating the individual based model associated with a generalized logistic equation as given by Eq. 8 and how this yields exactly the same equations for adaptive dynamics as the standard logistic model. We also discuss how the conditions for evolutionary branching for both algorithms remain same. Lastly, we discussed how to simulate the birth death process when both birth and death events have a competition kernel associated with them while, at the same time, keeping the total population size bounded despite of evolutionary branching. Our analysis gives an intuitive interpretation of the canonical equation of adaptive dynamics as the translation of the population cloud in the phenotypic space due to birth-death events. Such an approach has interesting connections to the stochastic quantization in quantum field theory and the birth death events can be viewed as creation and annihilation operators of a quantum field [36]. A complete analysis of these connections is the focus of our future work.

Our simulations also enables us to address important questions regarding the nature of evolutionary dynamics. In [34] it was shown that phenotypic properties can combine in complicated ways and as a result evolutionary trajectories can exhibit chaos. It shows that, contrary to the traditional view, evolution is not a directional optimization process that goes to equilibration. Evolution might be a complicated trajectory in

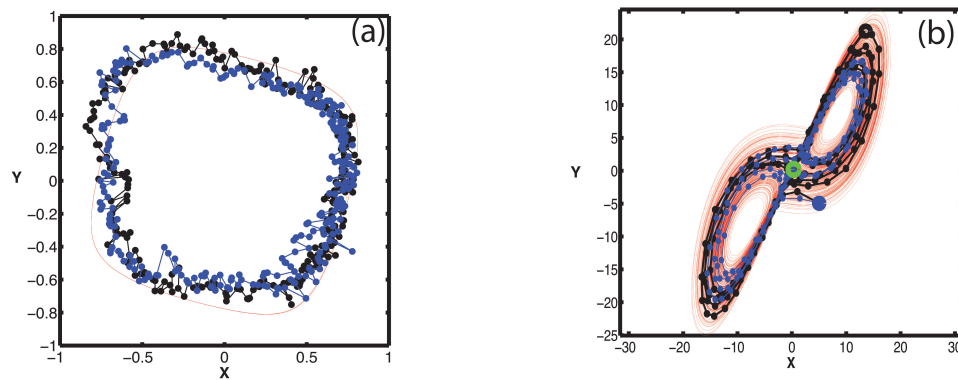


FIG. 2: Examples of correspondence between the monomorphic adaptive dynamics and the individual based model with the canonical competition kernel. (a) When the adaptive dynamics exhibits a limit cycle (red). Individual based models (black and blue) follow the monomorphic trajectory with stochastic fluctuations due to a small population size (about 100 individuals). (b) When the adaptive dynamics is the Lorenz attractor (red). Individual based models (black and blue) show the “butterfly effect”. Starting from *exactly same* initial conditions (the green dot), they evolve to different locations on the attractor (the black and the blue dot). (c) Example of another chaotic adaptive dynamics (red) and correspondence with the individual based models (black and blue).

the phenotypic space that is continuously twisting, turning and folding upon itself on a chaotic attractor. Our numerical simulations involving individual based simulations confirm that chaotic trajectories of evolution are sensitive to very small mutational changes. Even when these trajectories start with exactly the same initial conditions, they can end up in completely different regions of the phenotypic space. Therefore, if Gould’s tape of life were to be run again, with exactly same initial conditions and dynamics, we can end up with com-

pletely different biodiversity on earth. Moreover, the size of the mutations to cause this unpredictability of evolution need not be big as the underlying chaotic dynamics amplify small errors arising due to mutations. We hope our work is useful for simulations involving birth death processes in high dimensional phenotypic spaces as well as the study of evolution of language, culture, religions and agent based models in economics.

- 
- [1] M. Doebeli, “Adaptive diversification”, Princeton: Princeton University Press, (2011).
  - [2] U. Dieckmann and M. Doebeli, “On the origin of species by sympatric speciation” *Nature*, 400 (6742), 354357, (1999).
  - [3] O. Diekmann, “A beginners guide to adaptive dynamics”, *Banach Cent. Publ.*, **63**, 4786, (2004).
  - [4] S.A.H. Geritz, E. Kisdi, G. Meszna, J.A.J. Metz, “Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree”, *Evol. Ecol.*, **12** (1), 3557, (1998).
  - [5] Ulf Dieckmann and Richard Law, “The dynamical theory of coevolution: A derivation from stochastic ecological processes”, *J. Math. Biol.*, **34**: 579-612, (1996).
  - [6] N. Champagnat, R. Ferriere, S. Meleard, “Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models”, *Theor. Popul. Biol.*, **69**, 297321, (2006)
  - [7] N. Champagnat, R. Ferriere, S. Meleard, “From individual stochastic processes to macroscopic models in adaptive evolution”, *Stoch. Models* **24**, Suppl. 1, 2-44 (2008).
  - [8] D.T. Gillespie, “A general method for numerically simulating the stochastic time evolution of coupled chemical reactions”, *J. Comput. Phys.*, **42** (4) 403434, (1976).
  - [9] R. M. May *Nature* **261**, 459 467, (1976)
  - [10] J.A.J. Metz, R.M. Nisbet and S.A.H. Geritz, “How should we define fitness for general ecological scenarios” *Trends Ecol. Evol.*, **7** (6), 198202, (1992).
  - [11] S. Gavrillets, “Fitness Landscapes and the Origin of Species”, Princeton: Princeton University Press (2004).
  - [12] M. Barluenga, K.N. Stolting, W. Salzburger, M. Muschick, A. Meyer, “Sympatric speciation in Nicaraguan crater lake cichlid fish”, *Nature*, **439**, 719723, (2016).
  - [13] S.H. Berlocher, J.L. Feder, “Sympatric speciation in phytophagous insects: moving beyond controversy?”, *Ann. Rev. Entomol.*, **47**, 773815, (2002)
  - [14] Coyne, J.A., Orr, H.A., “Speciation. Sinauer Associates, Sunderland, Mass”, (2004)
  - [15] Dbarre, F., Nuismer, S.L., Doebeli, M., “Multidimensional (Co)evolutionary stability”, *Am. Nat.* 184 (2), (2014), 158-171.
  - [16] M. Doebeli, U. Dieckmann, “Evolutionary branching and sympatric speciation caused by different types of ecological interactions”, *Am. Nat.*, 156, S77S101, (2000).
  - [17] M. Doebeli and Y. Ispolatov, “Complexity and diversity”, *Science*, **328**, 493497, (2010)
  - [18] M.L. Friesen, G. Saxer, M. Travisano, M. Doebeli, “Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in *Escherichia coli*”, *Evolution*, 58 (2), 245260, (2004).
  - [19] R.T. Gilman, S.L. Nuismer, D.-C. Jhwueng, “Coevolution in multidimensional trait space favours escape from parasites and



- pathogens", *Nature*, **483** 328330, (2012).
- [20] M.D. Herron and M. Doebeli, "Parallel evolutionary dynamics of adaptive diversification in *Escherichia coli*", *PLoS Biol.*, **11** (2), 1001490, (2013).
- [21] H.C. Ito and U. Dieckmann, "A new mechanism for recurrent adaptive radiations", *Am. Nat.*, **170** (4), E96E111, (2007).
- [22] H.C. Ito and U. Dieckmann, "Evolutionary branching under slow directional evolution", *J. Theor. Biol.*, **360**, 290314, (2014).
- [23] G.M. Le, J. Pluain, T. Hindr, R. Lenski, D. Schneider, "Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with *Escherichia coli*", *Proc. Natl. Acad. Sci. U S A*, **109**, 94879492, (2012).
- [24] O. Leimar, "Multidimensional convergence stability", *Evol. Ecol. Res.*, **11**, 191208, (2009).
- [25] O. Leimar, A. Sasaki, M. Doebeli and U. Dieckmann, "Limiting similarity species packing, and the shape of competition kernels", *J. Theor. Biol.*, **339**, 313, (2013).
- [26] J. Pluain, T. Hindr, G.M. Le, O. Tenaillon, S. Cruveiller, C. Mdigue, N. Leiby, W. Harcombe, C. Marx, R. Lenski, D. Schneider, "Epistasis and allele specificity in the emergence of a stable polymorphism in *Escherichia coli*", *Science*, **343**, 13661369, (2014).
- [27] P.B. Rainey and M. Travisano, "Adaptive radiation in a heterogeneous environment", *Nature*, **394** (6688) 6972, (1998).
- [28] R.F. Rosenzweig, R.R. Sharp, D.S. Treves and J. Adams, "Microbial evolution in a simple unstructured environment genetic differentiation in *Escherichia coli* Genetics", **137** (4), 903917, (1994).
- [29] D.E. Rozen and R.E. Lenski, "Long-term experimental evolution in *Escherichia coli*. VIII. Dynamics of a balanced polymorphism", *Am. Nat.*, **155**(1) , 2435, (2000).
- [30] P.G. Ryan, P. Bloomer, C.L. Moloney, T.J. Grant and W. Delport Ecological speciation in South Atlantic Island finches *Science*, **315** (5817) 14201423, (2007).
- [31] V. Savolainen, M.-C. Anstett, C. Lexer, I. Hutton, J.J. Clarkson, M.V. Norup, M.P. Powell, D. Springate, N. Salamin and W.J. Baker, "Sympatric speciation in palms on an oceanic island", *Nature*, **441** (7090), 210213, (2006).
- [32] H. Svardal, C. Rueffler and M. Doebeli, "Organismal complexity and the potential for evolutionary diversification", *Evolution*, **68** (11), 32483259, (2014).
- [33] I. Ispolatov, V. Madhok and M. Doebeli, "Individual-Based models for adaptive diversification in high-dimensional phenotype spaces", *Journal of Theoretical Biology*, **390**, 97-105, (2016).
- [34] I. Ispolatov and M. Doebeli, "Chaos and unpredictability in evolution", *Evolution* **68**, Issue 5 1365-1373, (2014).
- [35] I. Ispolatov, V. Madhok, S. Allende and M. Doebeli "Chaos in High Dimensional Dynamical systems", *Nature Scientific Reports* **5**, Article number: 12506.
- [36] G. Bianconi and C. Rahmede, "Unified framework for quasispecies evolution and stochastic quantization", *Phys. Rev. E*, **83**, 056104, (2011).